AD 701 134

DEVELOPMENT OF IMPROVED APTITUDE AREA COMPOSITES FOR ENLISTED CLASSIFICATION

Milton H. Maier, et al

Army Behavioral Science Research Laboratory
Arlington, Virginia

September 1969

Technical Research Report 1159

AD

AD701134

DEVELOPMENT OF IMPROVED APTITUDE AREA COMPOSITES FOR ENLISTED CLASSIFICATION

Milton H. Maier and Edmund F. Fuchs

MILITARY SELECTION RESEARCH DIVISION

BESRL

U. S. Army
Behavioral Science Research Laboratory

September 1969

This document has been approved for public release and sale; its distribution is unlimited.

49

BEHAVIORAL SCIENCE RESEARCH LABORATORY

An activity of the Chief, Research and Development

J. E. UHLANER
Director

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by BESRL. Please address correspondence concerning distribution of reports to: U. S. Army Behavioral Science Research Laboratory, Attn: CRDBSRL, Room 239, Commonwealth Building, 1320 Wilson Blvd., Arlington, Virginia 22209.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Behavioral Science Research Laboratory.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Technical Research Report 1159

AD

DEVELOPMENT OF IMPROVED APTITUDE AREA COMPOSITES FOR ENLISTED CLASSIFICATION

Milton H. Maier and Edmund F. Fuchs

MILITARY SELECTION RESEARCH DIVISION
Edmund F. Fuchs, Chief

U. S. ARMY BEHAVIORAL SCIENCE RESEARCH LABORATORY

Office, Chief of Research and Development
Department of the Army

Room 239, The Commonwealth Building
1320 Wilson Boulevard, Arlington, Virginia 22209

September 1969

Army Project Number
2Q062106A722

Differential Classification b-00

BESRL Technical Research Reports and Technical Research Notes are intended for sponsors of R&D tasks and other research and military agencies. Any findings ready for implementation at the time of publication are presented in the latter part of the Brief. Upon completion of a major phase of the task, formal recommendations for official action normally are conveyed to appropriate military agencies by briefing or Disposition Form.

FOREWORD

The DIFFERENTIAL CLASSIFICATION Work Unit applies psychological measurement methods to enable the Army to make best use of the different skills and aptitudes of its enlisted personnel through increasingly accurate and differentiated measures of individual potential. Research is conducted to maintain and improve the effectiveness of the Army Classification Battery and related techniques and of conditions which may interact with the classification tests and thus affect the basis for utilization of the enlisted input --changes in training programs and job content and environment, for example.

Nearing completion is a large-scale validation study of operational and experimental measures conducted across a full range of the Army's Military Occupational Specialties (MOS). The measures are evaluated for their effectiveness in predicting final grades in Army school training courses and subsequent performance in Army job assignments. Test and criterion data are the source of two major developments in the initial classification of enlisted men: 1) New and improved aptitude area composites based on the current operational Army Classification Battery (ACB) tests were developed, and 2) a long-term study has been conducted of the differential validity of both operational and experimental measures with a view to possible major reconstitution of the aptitude area system. The present report is an account of the first of these developments. A publication dealing with the long-term study is in preparation.

The entire research task is responsive to special requirements of the Deputy Chief of Staff for Personnel and the U. S. Continental Army Command as well as to objectives of RDT&E Project 2Q062106A722, "Selection and Behavioral Evaluation," FY 1969 Work Program.

J. E. UHLANER, Director
U. S. Army Behavioral Science
Research Laboratory

DEVELOPMENT OF IMPROVED APTITUDE AREA COMPOSITES FOR ENLISTED CLASSIFICATION

BRIEF

Requirement:

To develop new and improved aptitude area composites based on Army Classification Battery test scores for use in determining assignment of enlisted input to training courses.

Procedures:

In a longitudinal study of the ACB, operational test scores were obtained for 25,000 men in over 100 different Military Occupational Specialties (MOS). The men were followed through training and their course grades, including failing grades, served as criterion to determine how well the tests measured potential for training in each area. For combat MOS, course grades were supplemented by ratings on later performance in order to reflect some elements of the combat assignment. Based on the validity coefficients obtained, new aptitude area composites were developed through test selection methods. Simulated allocation studies were conducted to find how much the new composites would improve classification of men for training and jobs.

Findings:

The resulting eight aptitude areas can provide improved measured of trainability. The composites are each based on three or more tests. Thus, they would make greater use of information--already available from ACB test scores--about the potential of the individual for different occupational areas and permit increased differentiation in assignment. This in spite of the fact that tests of general ability are more heavily weighted than in the two-test composites.

Utilization of Findings:

With the revised aptitude area system, failure in Army training schools would be reduced by about 20 percent with concomitant savings in cost and fuller benefit from Army training resources.

Men who complete training would achieve better. The number of enlisted personnel performing at marginal level would decrease by about 10 percent and an increased number would perform at superior levels.

CONTENTS

PRECEDING PAGE BLANK

TABLES Page

FIGURES

DEVELOPMENT OF IMPROVED APTITUDE AREA COMPOSITES FOR ENLISTED CLASSIFICATION

Matching the abilities of men entering the Army to the demands of training courses and jobs is an ever-critical problem facing Army personnel management. Aptitude area composites have been developed to measure the potential of the men, and the scores from these composites are used in making decisions about the men's assignments. The productivity of the men in their training and job assignments can be enhanced by obtaining more accurate assessments of their potential to perform in the different job categories. Based on a longitudinal study of the effectiveness of the Army Classification Battery (ACB) tests in predicting training success in more than 100 Military Occupational Specialty (MOS) courses, an improved set of aptitude area composites has now been developed. These composites and the benefits that would derive from their use--specifically, how they could improve the productivity of the enlisted men in their training assignments--are described in the present report.

BACKGROUND

The Army Classification Battery and aptitude area composites were first introduced for operational use in 1949. A continuous research program has been in effect since then to maintain and improve the effectiveness of the ACB and the aptitude areas. As a result of research conducted in the early 1950's, a revised set of aptitude area composites was introduced operationally in 1956. A second major revision came two years later when two tests to predict combat performance were developed through a series of studies of performance of men in combat in Korea, on maneuvers in Germany, and under arduous climatic conditions in Alaska. The two tests were introduced operationally in 1958.[1]

The tests in the ACB and the operational aptitude area composites developed in 1958 and still used in 1969 are shown in Table 1. These composites will be referred to as the "1969 aptitude areas" to differentiate them from those developed in the research study reported here. Each test in the ACB measures a different facet of a man's potential. The Verbal and Arithmetic Reasoning tests are measures of general learning ability; these aptitudes are highly related to academic success. The Mechanical Aptitude Test is a general measure of mechanical ability, and the Pattern Analysis Test is a general measure of ability to visualize spatial relationships. The remaining tests measure more specialized

---

[1] Descriptions of earlier research on the ACB and composites are contained in BESRL Technical Research Reports 1095 (Reconstitution of the aptitude areas, 1956) and 1110 (Development of combat aptitude areas, 1958).

aptitudes. The Army Clerical Speed Test and the Army Radio Code Aptitude Test measure perceptual speed and accuracy; the clerical test deals with visual perception and the radio code test with auditory perception. These two tests plus the Pattern Analysis Test are the only ones in the Army Classification Battery that do not involve reading. All the other tests in the battery require the ability to read and comprehend an item. The Shop Mechanics, Automotive Information, and Electronics Information tests are useful measures of interest and aptitude, not at the level of the experienced worker in such jobs but at the level of the young men who can learn to do such jobs. The General Information Test is a complex test that spans both the general learning ability and mechanical areas.

The final test, the Classification Inventory, is unique in the battery. Whereas the other tests require the examinee to perform tasks that have a right answer, the items in the Classification Inventory ask the man to describe himself--what he has done, what he likes to do, and how he sees himself as a leader. The items in this test were found to be valid predictors of combat success in research conducted in Korea during active conflict. In research currently under way, improved predictors of combat effectiveness as evaluated in the Vietnam war are being developed.

Each of these tests was included in the battery because over the years BESRL research has demonstrated that it is an accurate predictor of success in one or more occupational areas. However, the demands of the training courses are more complex than the kinds of potential measured in any one test. A better job of predicting success can be obtained if the tests are combined into aptitude areas, shown in the right-hand portion of Table 1.

As reconstituted in 1958, the aptitude area system utilized eight composites of ACB tests; these composites are related to the occupational areas of the Military Occupational Specialty (MOS) structure. The occupational areas served by the aptitude areas are indicated by the titles of the aptitude areas. Each composite contains only two tests because these combinations of tests were found to be as satisfactory as more complex composites for predicting training success. In each area, one test measures general ability and one measures a specialized aptitude.

The Research Study

During the past ten years, Army equipment has become more complex, and training has had to keep pace. Hence, some schools have asked for more comprehensive measures of trainability. In response, BESRL has just completed a long-term research study to determine how the ACB test scores could be combined to provide better measures of trainability for the modern training courses. The study included about 25,000 men in over 100 different training courses representative of the training courses open to new enlisted men. The 25,000 men came from all over the country and were representative of input to each of the courses. Each of the occupational areas was well represented in the analysis. The sampling of men and training courses provided a scientifically sound base for developing improved aptitude area composites.

Table 1

THE ARMY CLASSIFICATION BATTERY AND THE APTITUDE AREA COMPOSITES IN USE IN 1969

ARMY CLASSIFICATION BATTERY		OPERATIONAL APTITUDE AREA COMPOSITES		
Test	SYMBOL	Title	SYMBOL	Formula
Verbal	VE	Infantry - Combat	IN	$\frac{AR + 2CI}{3}$
Arithmetic Reasoning	AR			
Mechanical Aptitude	MA	Armor, Artillery Engineers - Combat	AE	$\frac{GIT + AI}{2}$
Pattern Analysis	PA	Electronics	EL	$\frac{MA + 2ELI}{3}$
Army Clerical Speed	ACS			
Army Radio Code	ARC	General Maintenance	GM	$\frac{PA + 2SM}{3}$
Shop Mechanics	SM			
Automotive Information	AI	Motor Maintenance	MM	$\frac{MA + 2AI}{3}$
Electronics Information	ELI	Clerical	CL	$\frac{VE + ACS}{2}$
General Information	GIT			
Classification Inventory	CI	General Technical	GT	$\frac{VE + AR}{2}$
		Radio Code	RC	$\frac{VE + ARC}{2}$

The study design was a longitudinal one in which the operational ACB scores were obtained for the men before they started their training courses. The men were then followed through training and their course grades were obtained. The new composites were developed by finding the combinations of tests that were the best predictors of success in training and still were simple enough for use in the field. Since the men had already been assigned to their training courses, a statistical correction was required to make the results generalizable to the full mobilization population. The composites shown in Table 2 were selected as the combinations of tests that provide the most satisfactory measures of trainability.

In the new aptitude area system, each area consists of three or four tests instead of only two. The existing relationship between aptitude areas and the MOS structure was retained. Each area contains the Arithmetic Reasoning Test, and six of the eight areas contain either the Verbal or General Information test. Thus, a measure of general learning ability is present in each and every composite.

In the proposed composites, as in the 1969 system, each test is weighted according to the unique contribution it makes to the prediction of training success. The weights are as shown in Table 2. For example, in computing the Infantry (IN) Aptitude Area, the Arithmetic Reasoning and Shop Mechanics tests each have a weight of 1, while the Classification Inventory has a weight of 2. This computing formula means that the Classification Inventory is twice as important as the other tests in the prediction of success in Infantry jobs. The weighted test scores are added together; the sum is divided by the appropriate divisor and a constant is subtracted to convert the scores to the Army standard score scale, which has a mean of 100 and a standard deviation of 20. The reason for placing the aptitude area scores on the Army standard scale is to make all the scores comparable; for example, in all areas a score of 90 means that 30 percent of the mobilization would score below that point.

New Composites

With the more accurate measurement provided by the new composites, the Army training resources can be used more productively. The first benefit that could accrue is that the number of training failures would be reduced. In 1967, there were about 25,000 failures. With the proposed system, the number could be reduced by about 20%, or to 20,000. According to the Army Comptroller's office, the average cost of putting a man through advanced individual training is $2,000. For the 5,000 failures, at $2,000 each, a gross of $10,000,000 was largely wasted, so far as utility to the Army is concerned, in attempting to train these men. By using the proposed aptitude area system, $10,000,000 of training funds per year could be used productively rather than non-productively. It is emphasized that the selection standards would not have to be raised to realize these gains; the benefits would accrue from assigning the men to the training courses that are better suited to their abilities.

Table 2

PROPOSED APTITUDE AREA COMPOSITES

Aptitude Area Composite	Add: Army Standard Scores for Component tests	Divide by[a]	Subtract[b]
IN	AR+SM+2CI	3	33
AE	2AR+3ACS+3AI+2GIT	7	43
EL	2AR+PA+MA+2ELI	5	20
GM	AR+MA+SM+GIT	3	33
MM	AR+2AI+ELI+GIT	4	25
CL	2AR+2VE+ACS	4	25
GT	3AR+PA+ACS+2GIT	6	17
RC	3AR+2PA+4ARC+2GIT	10	10
GA[c]	AR+VE	2	--

[a]The divisors are smaller than the sum of the weights to compensate for the reduction in standard deviations resulting from the differences from unity in test intercorrelations.

[b]The constant must be subtracted to adjust the mean score for the difference between sum of weights and divisor.

[c]GA stands for General Aptitude. It will be used for general screening in the same way that GT is used under the current system, as in deciding whether the enlisted man should be given the Officer Candidate Test or Army Language Aptitude Test.

The second benefit is that men who complete the course would perform at a higher level in the training course. To determine how much improvement would be obtained, the current enlisted input was simulated on the computer, and the simulated individuals were assigned on the basis of their aptitude scores under both the 1969 and proposed sets of scores. The expected level of training success was evaluated for each set of composites. Men at two levels of expected performance were of special interest: the marginal man, who performs at a level comparable to 90 on the Army standard score scale, and the superior man, who performs at a level comparable to 120 on the Army standard score scale. As expected, the number of men at the marginal level decreased and the number of men at the superior level increased when assigned by the proposed composites. Using the 1969 composites, with a total annual input of 500,000, there would be about 113,000 marginal performers in training. Using the proposed composites, the number of men performing at marginal level would drop by 12,000 or by about 10 percent. Some men who would have

been marginal in an inappropriate assignment would be average in a different area. These improvements can be realized without changing the selection standards.

In sum, through the more accurate measurement of aptitudes, the training resources of the Army would be used more effectively. The number of failures would be reduced, resulting in saving of time and money. The time and money saved could be spent more productively on those who complete training. Because the assignments would be more appropriate, the men could learn more and the instructors could do a better job.

In the discussion of the benefits that the Army would realize from operational implementation of the proposed composites, the gains that could be obtained from improved selection were not considered. If the proposed composites were to be used for selection at the AFEES, the selection standards would in effect be raised. The reasons lie in the selection standards in effect in 1969 and in the meaning of a qualifying score of 90. Under these mental standards, Category IV (AFQT percentile scores of 10-30) non-high school graduates have to achieve one or two qualifying scores of 90 or better. These individuals can achieve qualifying scores of 90 more easily under the 1969 aptitude area system than they could under the revised system. This situation came about because the requirement tends to capitalize on the individual's two highest test scores. The rest of his test scores, which must be lower, have been ignored. If some of the lower test scores were included in computing his aptitude scores, his highest aptitude area score could often be expected to drop below 90. The lower aptitude area scores are more accurate estimates of his trainability, and men who do not attain the required scores should, in fact, be considered unqualified.

In summary, a new set of aptitude area composites has been developed for use in the Army personnel system. The new composites provide a more comprehensive measurement of trainability in each area, with a resulting improvement in the accuracy of predicting training success. Considerable improvement in the training assignments can be realized through use of the new composites; the benefits would be about a 20% reduction in the number of training failures, and about a 10% improvement in the training performance of the men completing training successfully.

TECHNICAL SUPPLEMENT

OVERVIEW

In a longitudinal study of the effectiveness of the ACB tests as predictors of success in training, operational test scores were obtained for about 25,000 enlisted men in over 100 different MOS courses. The men were followed through their training to find out how successful they were. Their course grades, including failing grades, served as the criterion to determine the effectiveness with which the ACB tests measured training potential for each of the training areas. On the basis of the validity coefficients obtained, new composites were developed to provide better measures of trainability. Simulated allocation studies were conducted to find how much the new composites would improve the training assignments of men coming into the Army.

Details of the data collection, the statistical analysis, and the results are presented in the following sections.

METHOD

Sampling and Data Collection

The number of MOS studied in each occupational area is shown in Table 3. Only three MOS are shown for the Infantry (IN) area and two for the Radio Code (RC) area because these were the only courses open to new accessions in the areas. In the rest of the areas, the sampling of courses was sufficiently comprehensive to provide stable estimates of the validity of the ACB tests. The complete listing of MOS courses is shown in Table A-1 of the appendix.

In each sample, successive classes were included until the desired number of cases was available. The number of usable cases is also shown in Table A-1. For courses taught at more than one installation, several installations were included. For a course taught at only one installation, the sampling necessarily was limited to that one place. Data were collected for all persons entering the courses during the period of data collection, from late 1964 through the middle of 1965. The analyses were restricted to males who had no prior Army work experience. If an individual had an MOS indicating that he had performed in an Army duty assignment, he was dropped from the study. Some entire MOS samples were dropped from the study because the bulk of the input had had prior Army experience, and not enough recent accessions remained for study; these samples are indicated in Table A-1. Enlisted women were not included. Reservists and National Guard members on active duty for training were retained in the study. The individuals retained in each sample were similar to new recruits coming into Army who receive their initial training assignments largely on the basis of their aptitude area scores.

Table 3

NUMBER OF MOS FROM EACH OCCUPATIONAL AREA
SAMPLED IN ACB LONGITUDINAL STUDY

Area	N
Tactical Operations (IN)	3
Tactical Operations (AE)	13
Missile and Fire Control Electronic Maintenance (EL)	16
General Electronic Maintenance (EL)	12
Precision Maintenance (GM)	8
Auxilliary Services (GM)	8
Motors (MM)	15
Clerical (CL)	19
Graphics (GT)	5
General Technical (GT)	12
Radio Code (RC)	2

Variables

The variables were the 11 tests from Army Classification Battery or the Army Qualification Battery. The tests in the Army Qualification Battery (AQB) are parallel to counterpart forms in the ACB (1) and no operational distinction is made between the scores from the two batteries. For ease of discussion, reference is made only to the ACB, although some scores were also obtained with the AQB. Scores on an experimental battery were also obtained. Results are not included in the present report but will be presented in a later publication.

The criteria were the course grades obtained in training. Ratings of on-the-job performance were also obtained, but analysis of these will also be presented later. Course grades as reported by the training facility were accepted as indicators of the knowledge and skills the graduates had upon completion of their training. For the IN and AE areas, course grades were supplemented by ratings of performance on the job. The job ratings contain evaluation elements that more accurately reflect performance under stress of combat. Data were available from a large-scale study of men in the combat areas who were evaluated

during maneuvers in CONUS and in Europe (2). Validity coefficients from this study were averaged in with the training grades to produce the final criterion vector for the IN and AE areas.

A large proportion of the trainees, about 25%, did not graduate on schedule, and the course grades for this group, if available, were of questionable meaning. The exceptional cases consisted of three main subgroups: academic failures, academic turnbacks or recycles, and withdrawals for non-academic reasons. The withdrawals were assumed to be randomly distributed, and were dropped from the study. A separate analysis was conducted to determine how to assign criterion grades to the failures and turnbacks (3). The failures and turnbacks were not randomly distributed and had they been dropped, the samples would have had restricted distributions of criterion scores. The procedure for determining the grades to assign to failures and turnbacks was to find the regression of grades on a predictor of known validity, and then to place the failures and turnbacks on this line. Results indicated that this could be achieved by assigning a score one standard deviation below the minimum passing grade to the failures. For the turnbacks, two procedures were followed: If a final grade was not reported, the grade one-half of a standard deviation below the mean of the graduates was assigned; if a final grade was reported, that value was included in the analysis. With the inclusion of the failures and turnbacks, the samples had no known source of criterion bias.

RESULTS

Validity Analysis of ACB tests

The validity of the 11 ACB tests for predicting course grades was computed in each sample, yielding 110 validity vectors. Since the students in each class had already been selected on the ACB tests, the validity coefficients were subject to restriction in range. The multivariate correction for restriction in range was applied to each validity vector, so that all the samples were on a common basis. The correlation matrix and standard deviations used in the correction are shown in Table 4. The corrected coefficients are estimates of the validity that each test has in the full mobilization population.

Each validity vector was examined for reasonableness and for consistency with the values obtained for the other MOS in the same cluster. An MOS cluster consists of all MOS that have the same aptitude area as a selector; for example, MOS beginning with a 2 or 3 have EL as the selector, and are in the same cluster. Samples whose coefficients were uniformly near zero were discarded. Such low coefficients meant that the criterion grades were questionable. The ACB tests have a long history of being valid predictors and even in the present study the tests were valid for similar courses or identical courses taught at different installations. The samples that were deleted from the study because of low validity are indicated in Table A-1.

Table 4

INTERCORRELATIONS[a] OF TESTS OF THE ARMY CLASSIFICATION BATTERY

ACB Test	Intercorrelations[b] VE	AR	PA	MA	CI	ACS	ARC	GIT	SM	AI	ELI
Verbal (VE)	VE										
Arithmetic Reasoning (AR)	68	AR									
Pattern Analysis (PA)	49	55	PA								
Mechanical Aptitude (MA)	54	55	56	MA							
Classification Inventory (CI)	39	34	31	35	CI						
Army Clerical Speed (ACS)	43	48	35	32	27	ACS					
Army Radio Code (ARC)	47	48	41	41	27	42	ARC				
General Information Test (GIT)	65	57	46	58	41	36	41	GIT			
Shop Mechanics (SM)	40	38	50	59	31	22	27	51	SM		
Automotive Information (AI)	34	35	37	56	25	14	22	51	63	AI	
Electronics Information (ELI)	47	44	47	59	32	24	28	53	56	57	ELI
AFQT	72	73	72	67	39	40	47	64	60	52	58
SD	21.4	21.3	22.3	18.9	21.4	18.3	26.8	18.5	18.2	19.8	20.5

[a] Decimal points omitted.

[b] The correlation coefficients are based on about 26,500 cases tested during the spring of 1966.

Visual inspection did not reveal any MOS that were markedly deviant from the pattern of validity coefficients and that were similar to the pattern for another cluster. Only the pattern and not the level of the coefficients was considered. The operational eight clusters of MOS were confirmed by this procedure. The validity coefficients for the MOS in each cluster were averaged to obtain the best estimates of the validity of the ACB tests. The averages are shown in Table 5. The proposed composites were selected on the basis of these mean validity vectors.

Test Selection

A test selection was performed on each validity vector, following the procedure developed by Summerfield and Lubin (4). The purpose was to find that subset of tests which could best be combined into an aptitude area composite for use in the field. The multiple correlation of all 11 tests was computed to determine the maximum validity that could be obtained with the battery. The multiple correlations and beta weights are shown in Table A-2 of the appendix. Not all 11 tests could be used in the composites because computation would be too cumbersome in the field and also because some of the tests would have negative weights. Tests with negative weights were dropped from consideration because such weights lack face validity and would thus be undesirable from an operational point of view.

Since all the tests had positive validity for training performance, any negative weights arose because of the pattern of intercorrelations among the predictors and criterion rather than because of an intrinsic negative relationship between test score and training grades. In the absence of a negative validity coefficient, tests with negative weights would be difficult for a user in the field to interpret.

Two considerations entered into selection of the tests for the proposed composites. The first was that the multiple correlation coefficient of each composite should be within one or two points of the maximum attainable with the entire battery. The second was that the sum of the beta weights of the selected tests be large relative to the sum of the weights of all tests in the battery.

The sum of the weights in the composites must be relatively large if bias in the classification system is to be avoided. The bias would arise because some individuals would appear to be higher in one aptitude area, but with more complete measurement actually would be higher in another. The diagrams in Figure 1 may help to illustrate how all the tests in the battery influence classification and assignment even though they are not included in a composite.

Table 5

ACB VALIDITY COEFFICIENTS[a] FOR CLUSTERS OF MOS

Cluster	ACB Test										
	VE	AR	PA	MA	ACS	ARC	SM	AI	ELI	CI	GIT
IN	28	34	28	31	24	20	31	28	22	31	34
AE	30	36	29	33	32	23	30	33	24	23	34
EL	51	57	49	52	37	37	44	41	53	31	51
GM	52	53	45	53	31	37	49	46	47	30	54
MM	47	49	41	51	30	34	45	54	51	30	54
CL	61	61	41	42	44	41	26	28	37	29	50
GT	54	58	46	49	41	41	37	32	43	32	52
RC	45	48	40	42	30	46	33	29	35	22	43

[a]Decimals omitted from the correlations.

In diagram A of Figure 1, the relationship between the 1969 EL and GT aptitude area composites is shown; these two areas were shown for purposes of illustration. The following argument applies to all aptitude areas. The diagonal in diagram A shows the points at which the EL and GT scores are equal. Individuals above the diagonal have higher GT scores, and in the absence of quota restraints they would be assigned to GT jobs. Those below the diagonal have higher EL scores, and they would be assigned to EL jobs. With quotas, the principle would remain the same but assignments would be based on a more complex weighting scheme. Also, in the operational system, all eight aptitude area scores would have to be considered simultaneously.

In the two-test composites, the EL composite consists of MA plus 2 ELI, and the GT of VE plus AR. According to these composites, the AR score has no influence in the prediction of success in the EL jobs. In the development of the proposed composites, however, it was found that AR contributed to EL as well as to GT. If AR were added to the EL composite, then some of the individuals who previously scored higher on GT would now score higher on EL and some individuals would move in the opposite direction. The direction of the movement would depend on the magnitude of the AR score relative to the other test scores in the composites. The proposed EL composite results in more accurate classification because it includes AR, which makes an independent contribution to the prediction of training success in EL, as shown in Table A-2.

The relationship between pairs of the proposed composites is more like that shown in diagram B of Figure 1. The proposed composites are more highly intercorrelated than the two-test composites. Large differences between scores would therefore occur less frequently, and the scores would cluster around the diagonal line showing equal scores. Although the differences in diagram B are smaller on the average than in diagram A, those that do occur are more likely to reflect true differences in training performance. The classification and assignment system based on the proposed composites would result in fewer misclassifications because it is based on more of the valid information available from the entire battery.

The results of the test selection are shown in Table 6. The tests are listed in the order in which they were selected, except for the AE cluster. The first test listed in each area is the one that had the highest validity coefficient; the second test had the highest correlation with the criterion after the influence of the first test has been removed, and so on for the remaining tests. In the AE cluster, the GIT was forced as the first test to be selected because of its demonstrated validity against estimated performance during combat maneuvers (5). The multiple correlation coefficients for the composites are shown in the table along with the maximum validity obtainable for the entire battery. As can be seen, the coefficients for the composites were always within one or two correlation points of the maximum. The weights for all the tests in the full regression equations, presented in Table A-2 of the Appendix, show that in general the tests with the highest weights were the ones selected for the proposed composites.

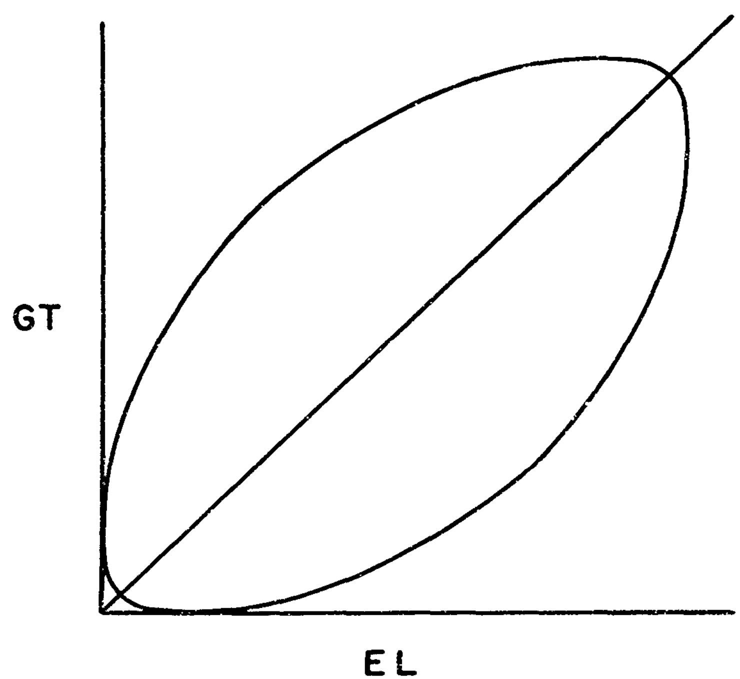

Diagram A: Composites used in 1969

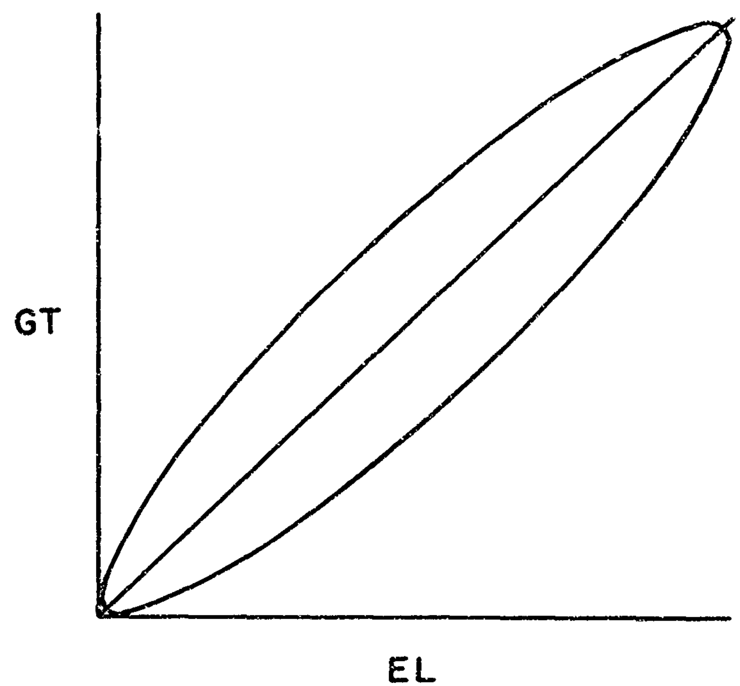

Diagram B: Proposed Composites

Figure 1. Illustration of Relationship Between EL and GT Composites

The beta weights, which show the relative importance of each test in the composites for predicting the criterion, and the integer weights proposed for operational use are also shown in Table 6. The proposed weights were roughly proportional to the beta weights for all tests except the Classification Inventory (CI) in the IN area and the General Information Test (GIT) in the AE area. According to the beta weights, the CI and GIT should have had a weight of one, but the expert judgment of the research staff based on prior studies was that these should be given a greater weight. Prior research studies to validate the CI against ratings of combat performance obtained during the Korean War demonstrated that CI was one of the most valid predictors for the more ultimate criterion, and hence it was given a higher weight (5). The weight for GIT was also raised because of its high validity against performance in combat maneuvers. The first set of proposed composites (Table 2) is consistent with earlier results obtained by Helme (6).

The proposed composites were divided by a constant other than the sum of the weights. The reason for using the divisors shown in Table 2 was to set the standard deviations approximately equal to 20. For example, if the sum of the weights had been used as the divisor for IN, the standard deviation would have been 14, but by dividing by three instead of four, the standard deviation was increased to 19. The mean aptitude area score had to be set at 100, which for IN was accomplished by subtracting 33. With these adjustments, the aptitude area scores were placed on the Army standard score scale, and all scores can be interpreted in the same way. A practical significance is that about 30% of the normal Army input should have scores below 90 in any aptitude area; if the standard deviations were less than 20, less than 30% would have scores below 90. A score of 90 is frequently used as a minimum prerequisite for Army training courses.

For operational purposes, an additional score would be used in the selection and classification procedure. This score has been designated General Aptitude (GA), and is identical to the GT aptitude area score in the current composites (the average of Verbal and Arithmetic Reasoning test scores). The GA score would be used to determine who is eligible to take additional tests with an academic orientation, such as officer candidate tests, flight training tests, and language aptitude tests. The GA score would not be used in making training assignments.

Cross Validation

The validity coefficients based on all the samples are the most stable estimates of the validity of the ACB that could be obtained with the data available. The proposed composites, then, include the tests that yield the best prediction of training success. As with all sampling, however, the problem of generalizability arises. The correlation values may be inflated because of capitalization on chance covariation in the test scores and the course grades. Further analyses were conducted to determine how well the results generalized from one set of data to another where there was no possibility of capitalizing on chance.

Table 6

RESULTS OF TEST SELECTION

MOS Cluster	Tests Selected			Validity Coefficients[a]	
	Test	Beta[a] Weight	Proposed Weight	Composite	Full ACB
IN	AR	2118	1	45	45
	CI	1846	2		
	SM	1723	1		
AE	ACS	1901	3	46	47
	AI	2108	3		
	AR	1503	2		
	GIT	0784	2		
EL	AR	3250	2	67	68
	ELI	2594	2		
	MA	1246	1		
	PA	1156	1		
GM	GIT	2002	1	65	67
	AR	2606	1		
	SM	1920	1		
	MA	1566	1		
MM	GIT	1997	1	66	67
	AI	2755	2		
	AR	2140	1		
	ELI	1486	1		
CL	AR	3148	2	68	69
	VE	3356	2		
	ACS	1454	1		
GT	AR	3164	3	65	66
	GIT	2276	2		
	PA	1432	1		
	ACS	1219	1		
RC	AR	2161	3	57	58
	ARC	2530	4		
	GIT	1504	2		
	PA	1091	2		

[a]Decimal points omitted from the Beta Weight and correlations.

The first step was to divide the MOS in each cluster into two halves, labeled A and B. The division was accomplished on the basis of MOS titles. In so far as possible, MOS that appeared to involve similar content and functions were placed in different halves; for example, if the same MOS was taught at two installations, one sample was placed in A, and the other in B. No division of the MOS in the IN and RC clusters was made because there were only two MOS in each cluster. For the other six clusters the MOS were divided, and the same analyses performed on the full cluster were also performed on each half. The validity coefficients were averaged, and test selections were performed on each validity vector. Each of these validity vectors and composites is subject to chance variations, just as those based on the full clusters. To avoid the chance factors, the composites derived in one subsample were applied to the MOS in the other. The validity coefficients crossed over to the other half-clusters are free of any inflation, and they provide unbiased estimates of the validity of the aptitude area scores.

The validity vectors of the ACB tests for each half of the MOS are shown in Table 7. The coefficients shown here are comparable to those shown in Table 5, except for the AE vectors. The AE validity vector in Table 5 included training grades and job ratings in the criterion, while that in Table 7 included only training grades. Only the vectors for the two CL halves showed close similarity; all the other areas in Table 7 showed a difference of at least five points for one or more tests.

The differences in the validity vectors bring out the need to include a large variety of MOS in each cluster to obtain composites that can be generalized. Each MOS has its unique profile of validity coefficients. By itself, this profile would ordinarily not generalize to other MOS. By combining a large number of MOS, however, the differences among the MOS tend to average out, and what is left is the general pattern of aptitudes common to all the MOS in the cluster.

Test selection was performed on each validity vector to determine whether the same abilities were required for the two halves of MOS in each cluster; the results are shown in Table 8. In general, the same abilities were included in the two halves, although there were exceptions. In the CL area, the tests were identical because the validity vectors were similar. At the other extreme were the EL and GT areas, which had two tests in common. In these areas, the other two tests for each half cluster were different--PA and GIT in EL-A vs ACS and AI in EL-B; MA and ACS in GT-A vs PA and ARC in GT-B. In the GM area, the two composites were similar in that two tests were identical (AR and GIT) and the third test was similar (AI and SM); only the last test was different (VE vs MA). In MM, three tests were the same, but the other test was different--MA in MM-A and VE in MM-B. The two composites in the AE area appear to measure similar aptitudes.

Table 7

ACB VALIDITY COEFFICIENTS[a] FOR HALF CLUSTERS OF MOS

MOS Half-Cluster	ACB Test VE	AR	PA	MA	ACS	ARC	SM	AI	ELI	CI	GIT
AE-A	45	46	35	38	32	32	27	36	36	22	43
AE-B	40	41	34	38	36	30	35	30	33	22	42
EL-A	52	56	50	52	35	36	44	39	51	31	52
EL-B	50	58	47	53	39	38	44	43	56	32	49
GM-A	56	56	47	55	32	39	50	49	52	33	57
GM-B	47	50	44	50	31	36	47	42	41	27	49
MM-A	53	55	47	57	34	38	50	54	56	30	59
MM-B	44	45	37	48	28	32	41	54	48	31	51
CL-A	62	61	39	42	44	40	27	29	36	29	50
CL-B	61	61	41	42	44	41	26	28	37	29	50
GT-A	50	52	43	48	39	38	34	31	41	31	50
GT-B	58	66	51	49	43	45	39	33	46	33	53

[a]Decimals omitted from the correlations.

Table 8

TEST SELECTION RESULTS IN HALF-CLUSTERS OF MOS

MOS Half-Cluster	Tests Selected and Integer Weights			
AE-A	2AR	2AI	2VE	1ACS
AE-B	1GIT	1ACS	1SM	1AR
EL-A	1AR	1ELI	1PA	1GIT
EL-B	3AR	3ELI	1ACS	1AI
GM-A	1GIT	1AR	1AI	1VE
GM-B	5AR	4SM	3GIT	3MA
MM-A	2MA	4AR	3AI	3ELI
MM-B	6AI	3AR	2ELI	2VE
CL-A	2VE	2AR	1ACS	
CL-B	2AR	2VE	1ACS	
GT-A	3AR	3GIT	3MA	2ACS
GT-B	3AR	1GIT	1PA	1ARC

The composites obtained on the full cluster, shown in Table 6, were similar to those found for the half-clusters. In the GM area, the tests of the proposed composites were the same as those in one of the half-clusters; in the EL, MM, and GT areas, features of each half were included. In the CL area, all composites were identical. The composite for the AE full cluster could not be compared directly with those in the half-clusters because the criteria were different; the criterion for the full cluster included both training grades and job ratings, while that for the half-clusters contained only training grades.

The unbiased validity coefficients for the aptitude areas, except for IN and RC, are shown in Table 9. The composites obtained in the A-half were validated against the MOS in the B-half, and vice versa. The average of the two coefficients, also shown in Table 9, provides an unbiased estimate of validity of the aptitude area scores in each cluster. The validity coefficients for the full clusters (proposed) are also shown for comparison. Generally, the unbiased coefficients were slightly lower than those for the proposed composites. The coefficients for the AE area could not be compared because of differences in the criteria. No unbiased estimate for the IN and RC areas could be obtained because there were not enough MOS to divide.

The results of the cross validation indicate that the proposed composites are reasonably suitable for all MOS in each cluster and that the validity coefficients can be generalized to other groups. There were enough differences in some of the clusters to raise questions about the homogeneity of the MOS. To determine whether the MOS clusters could be further reliably differentiated, additional analyses on the grouping of MOS are planned in conjunction with an analyses of experimental tests. The operational result of using more MOS clusters would be to make more effective assignments with a better utilization of the manpower resources.

SIMULATED ALLOCATION STUDIES

The proposed aptitude area composites provide more accurate measures of trainability than do the operational two-test composites. The question remained, however, of how the improved validity translates to improved assignments. To make such a determination, a series of allocation runs was made on a computer.

In the simulated allocation studies, a set of aptitude area scores for a sample of entities (simulated men) was generated by computer. The first step was to generate a vector of 11 random normal deviates for each entity. These deviates were then transformed through matrix multiplication to have the characteristics of ACB scores for a normal input population. The correlation matrix and standard deviations used to transform the normal deviates are those shown in Table 4. The simulated scores after the transformation were on the Army standard score scale, and they had the same intercorrelations as the actual ACB scores. The aptitude area scores were then calculated, and each entity was allocated to a

Table 9

VALIDITY COEFFICIENTS[a] OF COMPOSITES

Aptitude Area	Composite: 1969 Operational	Proposed	Cluster A[b]	Cluster B[b]	Average unbiased validity coefficient
IN[c]	37	41	--	--	--
AE[c]	39	46	49	49	49
EL	58	66	66	64	65
GM	55	65	59	68	63
MM	59	66	61	69	65
CL	62	68	68	68	68
GT	61	65	68	58	63
RC	53	57	--	--	--

[a] Decimal points omitted.

[b] Unbiased estimates of validity.

[c] The coefficients shown under the operational and proposed headings were obtained for a criterion which included training courses and job performance on the criterion, the coefficients under A and B included only performance in training courses.

craining opening. The allocation was done so as to maximize the expected performance of the total group, just as operational allocation does. The expected performance for each entity was then computed and the averages taken in each area. The computer program to accomplish the simulation studies has been described by Niehl and Sorenson (7).

The allocations and evaluations were performed several times, once for each set of aptitude area scores evaluated. The following composites were tried out:

1. Operational two-test composites

2. Proposed composites

3. The regression equation composites for all ACB tests

4. The A cluster composites

5. The B cluster composites

The same entities were allocated by each of these five methods; thus, any differences in the averages were a function of the composites and not of sampling differences in the entities. The effectiveness of each composite was evaluated by using the full regression equation to compute the expected performance of the entity in the area to which he was assigned. The beta weights used in the regression equations are shown in Table A-2 of the Appendix. The full regression equations were used to evaluate the composites in view of the analysis by Brogden (8) in which he demonstrated that the average expected performance based on the full equation approaches the average of the actual criterion performance as the number of men assigned increases.

A total of 30 samples with 300 entities in each was allocated by each of the six composites. Previous research by Sorenson (9) has shown that 300 entities in each sample provides stable estimates of the allocation average. The average expected performance in training across all 30 samples was obtained.

In the allocation runs, one restraint was that quotas be established for each aptitude area. The quotas were obtained by computing the percentage of men assigned to the MOS in each aptitude area during the period January through July 1967. The percentages were as follows:

IN	29%
AE	10
EL	10
GM	05
MM	18
CL	13
GT	11
RC	04

None of the other restraints in the operational allocation system, such as recommended MOS or transportation cost, were incorporated into the simulated runs. These restraints could not readily be built into the program; and in any case they do not affect average expected performance as much as does the quota restraint. The results of the simulated runs are shown in Table 10.

The figures shown in Table 10 are the mean expected performance for each method of allocation. The scale values are arbitrary and can be interpreted only in relation to each other. If the entities were assigned at random, that is, by ignoring the aptitude area scores, the mean expected performance would be 100. Since all the means in Table 10 are above 100, some improvement was realized by each of the methods. The smallest gain over random assignment (2.87 points) was realized with the two-test composites. The proposed composites resulted in a gain of 4.10 points, which is about a 40% gain over the operational composites. This increased gain would be achieved by using more of the information about the new recruits--information already available from the ACB. The maximum gain that could be realized from the ACB is 5.07 points, which came from the full regression equations.

Table 10

ALLOCATION EFFECTIVENESS OF ACB TEST COMPOSITES

Composites	Average expected performance
Two-test	102.87
Proposed	104.10
Full Regression Full Cluster	105.07
A Half-cluster	103.74
B Half-cluster	103.77

These results were as expected because of the way allocation works. In allocation, the assignments are made on the basis of aptitude area scores, which are estimates of trainability in the MOS clusters. The two-test composites make minimal use of the information about each man, and the estimates of trainability are not as accurate as those based on more tests. Each test in the battery contributes something to the accuracy of prediction, although after the third or fourth test has been included, the remaining tests add little accuracy. The big increase in the means is between the two-test composites and the proposed composites, which contain three or four tests each.

The reported gain for the full regression equations, 5.07 points, may be slightly biased because exactly the same equations were used to allocate the entities and then to compute expected performance. Any bias or inflation present in the allocation was also present in the evaluation, and there was no independent check on the effectiveness of the full regression equations. Although the true gain from using the full regression equations would still exceed that from the proposed composites, they are considered too complex to use and understand for routine operational use. They were included only to provide an upper bound for the gain that could be realized from using the ACB in allocation.

An independent check on the gain from the proposed composites is available from the A and B clusters. The composites used for these allocations were shown earlier in Table 8. The entities were allocated by the composites developed on the A half-clusters, and the expected performance was computed by full regression equations determined in the B half-clusters. A similar procedure was followed in evaluating B cluster composites. In this way, allocation and evaluation were based on completely independent samples, a procedure which provided a check on generalizability. Both these averages were smaller than that for the proposed composites, 103.74 and 103.77 vs 104.10. The expected performance from using the proposed composites would therefore not be much under 104.10 points, still a substantial improvement over the operational composites.

The results shown in Table 10 are averages across all the areas, and the effect of the proposed composites on each area was not considered. An important operational concern is that the composites should spread the available talent equitably across all areas. If one area receives a lesser proportion of high ability men, then not enough potential leaders and supervisors may be available for training in the MOS of that area. The averages in each area show how the allocation distributed the talent across the range of MOS. The results are shown in Table 11. The proposed composites resulted in higher averages in five areas, AE, EL, GM, MM, and CL. One area, IN, remained about the same, and two, GT and RC, dropped considerably. The big drop in GT, from 114.03 for the two-test composite to 99.27 for the proposed composite, arose because the verbal and reasoning aptitudes are spread out among all the areas in contrast to the two-test composites. Even though the average dropped in two areas, the overall effect was to raise the average expected performance.

The averages shown in Tables 10 and 11 are not directly comparable because the scale of measurement is different. Those in Table 10 were based on the full regression equations, which means that the standard deviation of the performance scores was a function of the validity of the aptitude area. Since the validity coefficients differed from cluster to cluster, predicted performance scores were on a different metric and could not be compared directly. This situation was corrected in Table 11 by arbitrarily setting all standard deviations of the predicted scores equal to 20. The effect of equating the standard deviations was to place the means on the Army standard score scale. The means, then, describe the entities assigned to each area in terms of their weighted ACB scores, corrected to full standard deviation, rather than in terms of predicted performance in the area of assignment. The effect was to increase the total allocation averages; for example, the average for the proposed composites increased from 104.10 in Table 10 to 106.31 in Table 11. Again these values are meaningless by themselves, and they must be interpreted relative to one another. The average for the proposed composites in Table 11 was 4% larger than the average for the operational composites. This increase was of the same order of magnitude as found when the full regression equations were used for evaluation.

Benefits to be Derived

One measure of improved utilization of available manpower is the rate of training failure. In FY 1968, the number of training failures was 23,300 out of a training base of about 500,000, a failure rate of 4.66%; this rate was obtained using the two-test composites, which had an average validity coefficient of .54. The average validity coefficient of the proposed composites was .59, an increase of .05. The increased validity, according to the Taylor-Russell Tables, would lower the attrition rate to 3.7% or by about 20%. With the same level of training input, the number of failures would be reduced by close to 5,000 men each year.

Table 11

ALLOCATION AVERAGES SHOWN BY APTITUDE AREA

Composite	Total	Aptitude Area							
		IN	AE	EL	GM	MM	CL	GT	RC
Operational	104.71	103.48	101.41	104.81	99.72	102.81	103.01	114.03	116.60
Proposed	106.31	103.49	108.28	106.17	109.84	105.82	115.38	99.27	109.79

The Taylor-Russell Tables are based on the normal bivariate surface. Determining that the failure rate would be reduced by 20%, it was assumed that the average proportion of successful trainees in an unselected group would be 80%, and that on the average the bottom 30% of the distribution would be screened out. With these assumptions, the failure rate is reduced from 5% to 4% as the validity coefficient increases from .55 to .60. The failure rate in all Army training courses during FY 1968 was slightly under 5%; the projected decline in failure rate can be expected to be an accurate forecast of experience under the proposed system.

Based on the results of the simulation studies, the proportion of men with predicted training performance at a level of 120 or higher and 90 or below was determined for each set of aptitude area scores; these levels are comparable to one standard deviation above the mean and one-half standard deviation below the mean, respectively. The results for each area and for the total are shown in Table 12. In all areas except GT and RC, the number of men with expected performance scores below 90, designated as marginal, was smaller for the proposed composites; the total percentage of marginal men was 20% for the proposed composites as compared to 23% for the operational composites. When these figures are translated into an annual input of 500,000 trainees, the 3% difference assumes more significance. Under the operational composites, about 115,000 men per year would be expected to be marginal in performance, whereas under the proposed composites the number would be about 100,000, a decrease of 15,000 marginal performers per year. On the superior side, those expected to perform at a level comparable to 120 or above, the proposed composites also were better. Based on an input of 500,000, about 110,000 men per year would be superior under the operational composites, and 120,000 men would be superior under the proposed composites. The improvement in expected performance arises through the improved validity and the resulting more appropriate assignments.

SCORE SCALE OF PROPOSED COMPOSITES

The focus of the present report has been to describe a new set of aptitude area composites which would improve the assignment and classification of Army enlisted men. The improvement in the new composites was realized because they capitalized on more of the valid information available from the ACB than did the two-test composites. With new composites of three or four tests each, more of the valid information is contained in each composite. At the same time, the change introduced problems about the units of measurement. Special adjustments were made in the proposed formulas to set all the standard deviations equal to 20.

The critical question in differential classification is whether an individual is higher in one aptitude area than in others. Whenever scores are compared, they must be on a common basis. In the operational Army system, aptitude area score differences have meaning only in terms of relative standing in the Army mobilization population, as indicated by the Army standard score. Thus, a score of 95 in IN is better than a score of 85 in MM because more persons in the reference group score below 95 than 85. If the higher scores did not imply higher relative standing, then the score differences would be exceedingly difficult to interpret.

Table 12

PERCENTAGE OF MEN GIVING MARGINAL AND SUPERIOR PERFORMANCES EXPECTED IN EACH APTITUDE AREA

Aptitude Area	Marginal Performance[a]		Superior Performance	
	Composite		Composite	
	Operational	Proposed	Operational	Proposed
IN	24%	24%	19%	20%
AE	29	16	16	27
EL	22	20	21	24
GM	33	16	15	31
MM	26	20	18	22
CL	23	08	18	39
GT	11	33	39	14
RC	05	16	42	30

[a]Marginal men are expected to perform at a level comparable to 90 on the Army standard score scale, and superior men at 120 on the same scale.

The tests of the ACB have all been placed on the Army standard score scale, which has a mean of 100 and a standard deviation of 20. Since all test scores have been equated in the mobilization population, the absolute values can be compared. The standard deviations shown in Table 4 are not exactly equal, but all except the Radio Code Aptitude Test are close enough to permit direct comparison. The proposed aptitude area composites, however, would not have had comparable standard deviations unless the adjustments in the divisor term had been made.

The classification and assignment system would be adversely affected if the standard deviations of the aptitude area scores were markedly unequal. The illustration in Figure 2 depicts what would happen if one area had a smaller standard deviation than another. For purposes of illustration, IN was assumed to have a standard deviation of 10 and EL one of 20. The ellipse shows the joint distribution of the scores in the two areas. The solid diagonal line at a 45-degree angle indicates equal absolute scores. Individuals above the diagonal would be classified as higher in EL and those below as higher in IN. Notice that this diagonal does not cut the ellipse down the middle. In the upper part of the curve, which contains the above average individuals, most of the cases are higher on EL, while in the lower part of the curve, which contains the below average individuals, most of the cases are higher on IN. This situation arises from the unequal standard deviations.

The broken diagonal divides the number of cases into two equal parts. The individuals in area b of the curve would be higher in EL in terms of difference in absolute scores, but they would be higher in IN in terms of relative standing in the reference population. The individuals in area d, on the other hand, would be higher in IN in terms of absolute values, but higher in EL in terms of relative standing. The individuals in area a would be classified in IN and those in area c to EL under both ways of computing the difference.

The operational allocation system would tend to assign the cases in area b to EL and those in area d to IN. Thus, the EL area would be favored with a greater number of above-average men, and the IN area would be depleted of men who were better in that area. One area would be favored merely because of a statistical artifact. If there is a requirement for more of the better men in a given area, the policy should be made explicit and the system adjusted to accomplish the goal.

The effects of assigning the men in area d in the illustration to IN would be most noticeable at the supervisory and leadership levels. The number of potential leaders would be smaller in IN, and if the discrepancy were great enough, the supply would not be adequate to meet the needs. In addition to the sheer proportion of potential leaders, the incremental value of adding high-quality men must be considered. As in the example, if EL already had a fair share of potential leaders, then adding more would not make a large contribution to overall performance; the effect might be compared to that of adding a pail of water to a pond. But if IN had a shortage of potential leaders, then adding more would have a greater impact on performance.

The need for equal standard deviations of the aptitude area score distributions may not be immediately apparent from an operational point of view. But, as was seen, the effects could be severe on the area that happened to have the smaller standard deviation. The answer to the question of whether an individual is higher in one area than another can be answered in an equitable manner only if the aptitude areas have a common score scale.

The question about differences in ability has been discussed in terms of aptitude area scores. The aptitude area scores are predictors of performance, and the predicted performance scores have the property that the standard deviation is a function of the validity; areas in which the validity is relatively low would have smaller standard deviations. If the assignments are to be made on the basis of predicted performance scores instead of aptitude area scores, then the effects would be the same as described above for different standard deviations of the aptitude areas; that is, the number and proportion of potential leaders would be lower in the areas with the lower validity. The operational system considers eight areas simultaneously, and the eight areas are positively intercorrelated. When a group of potential leaders is selected for one area, the supply of potential leaders for another area is depleted. The areas with the highest validity would be favored and those with lesser validity would suffer. The standard deviations should be equal whether validity is explicitly taken into account, as in the case of performance estimates, or whether the aptitude area scores themselves are used.

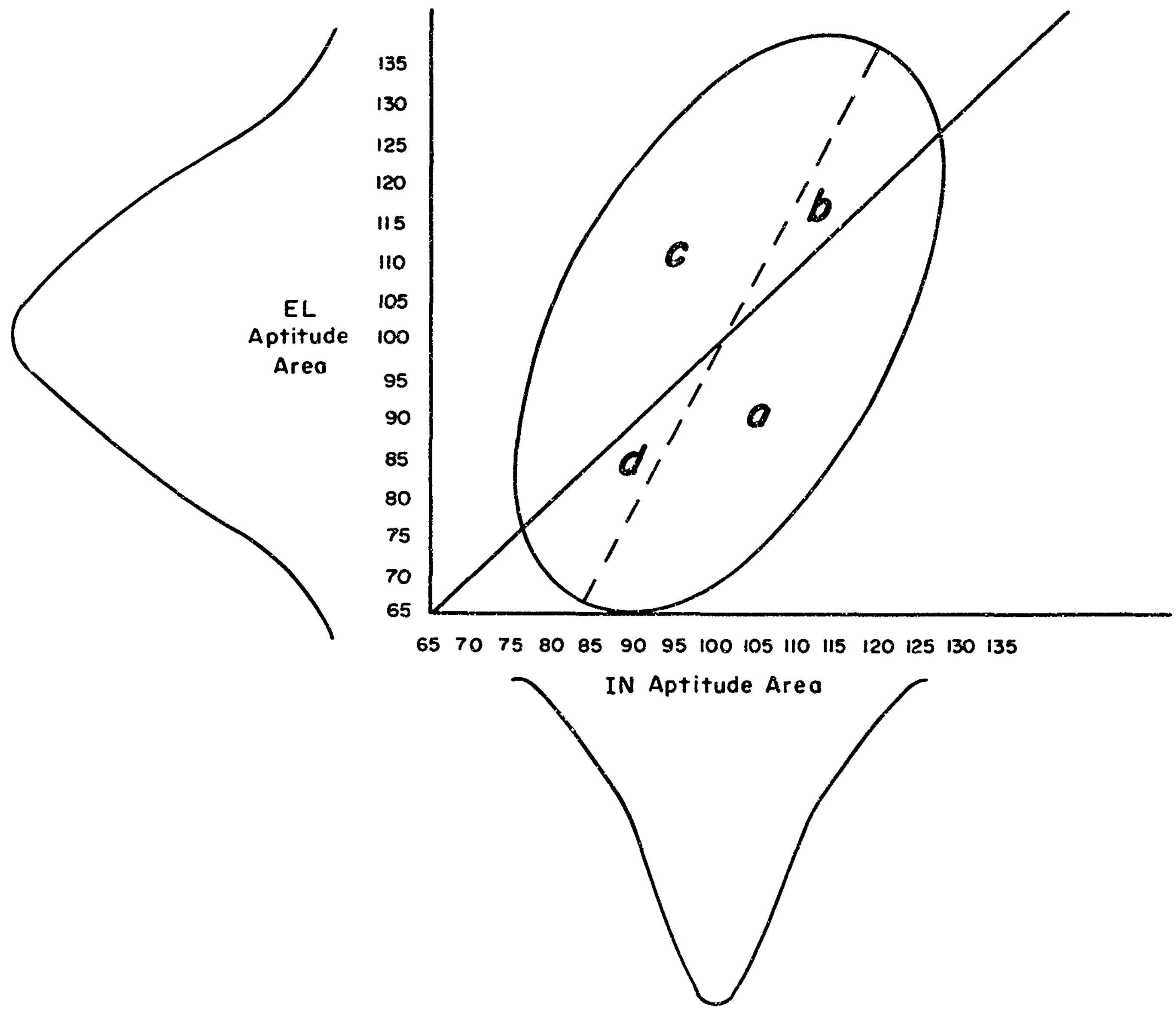

Figure 2. Illustration of Effect of Unequal Standard Deviations on Classification

LITERATURE CITED [2]

1. Seeley, L. C. and A. A. Anderson. Development of the Army Qualification Battery, Forms 2 and 3. BESRL Technical Research Report 1131. May 1963.

2. Willemin, L. P., A. H. Birnbaum, N. Rosenberg, and R. K. White. Validation of potential combat predictors in overseas maneuvers. BESRL Technical Research Note 80. July 1957.

3. Maier, Milton H. Procedures for assigning grades to failures and turnbacks in Army school courses. BESRL Technical Research Note 197. April 1968. (AD 833 582)

4. Summerfield, A. and A. Lubin. A square root method of selecting a minimum set of variables in multiple regression. Psychometrika, 1951, 16, 271-284.

5. Willemin, L. P. and E. K. Karcher, Jr. Development of combat aptitude areas. BESRL Technical Research Report 1110. January 1958. (AD 153 596)

6. Helme, William H. Evaluation of differential classification tests for the ACB. BESRL Technical Research Note 155. June 1965. (AD 621 698)

7. Niehl, Elizabeth and Richard C. Sorenson. SIMPO-I entity model for determining the qualitative input of personnel policies. BESRL Technical Research Note 193. January 1968. (AD 831 268)

8. Brogden, H. E. Least squares estimates and optimal classification. Psychometrika, 1955, 20, 249-252.

9. Sorenson, Richard C. Optimal allocation of enlisted men--Full regression equations vs aptitude area scores. BESRL Technical Research Note 163. November 1965. (AD 625 224)

[2] BESRL publications available from the U. S. Defense Documentation Center (DDC), formerly ASTIA, are indicated by inclusion of the assigned AD number in parentheses at the end of the citation. For publications not accessioned by DDC, the publishing agency or source is given. The U. S. Army Behavioral Science Research Laboratory (BESRL) was previously the U. S. Army Personnel Research Office (APRO).

APPENDIX

PRECEDING PAGE BLANK

Table A-1

MOS SAMPLED IN ACB VALIDITY STUDIES

Occupational Area and MOS Course		N[a]
TACTICAL OPERATIONS		
11B1	Light Weapons Infantryman	183 A
11C1	Infantry Indirect Fire Control	
11B1	Light Weapons Infantryman	197
11C1	Infantry Indirect Fire Control	
11B1	Light Weapons Infantryman	214
11C1	Infantry Indirect Fire Control	
11D1	Armor Intelligence Specialist	288
11E1	Armor Crewman	395
11H1	Infantry Direct Fire Crewman	98
12A1	Pioneer	417
13A1	Field Artillery, Basic	469
13E2	Field Artillery Operations and Intelligence Assistant	196
16B1	Hercules Missile Crewman	209
16C1	Hercules Missile Fire Control Crewman	149
16D1	Hawk Missile Crewman	205
16E1	Hawk Missile Fire Control Crewman	192 A
16H1	Air Defense Artillery Operations and Intelligence Assistant	97
17B1	Field Artillery Radar Crewman	171
17F1	Defense Acquisition & Surveillance Radar Crewman	60 A
17H2	Fire Distribution Systems Crewma..	92

Table A-1 (Continued)

Occupational Area and MOS Course		N[a]
<u>MISSILE AND FIRE CONTROL ELECTRONIC MAINTENANCE</u>		
21A1	Ballistic Missile Repair Apprentice	118
21E2	Sergeant Guidance Repairman	6 B
21H2	Pershing Guidance Control Repairman	6 B
21J2	Pershing Test Equipment Repairman	3 B
21K2	Pershing Computer Repairman	5 B
22A1	Electronic Repair Apprentice	205
22G2	Nike Launcher System Repairman	34 B
22J2	Hawk Missile Launcher Mechanic	174
22K2	Hawk Missile Launcher Repairman	63
22M2	Nike Missile Repairman	30 B
23S2 23T2	Hawk Pulse Radar Repairman Hawk Continuous Wave Radar Repairman	33 C
23P2	Hawk Fire Control Mechanic	113
23R2	Hawk CW Radar Mechanic	182
OF12	Fire Distribution Systems Electronics	242
26L20	Microwave Radio Repairman	180
<u>GENERAL ELECTRONIC MAINTENANCE</u>		
31E2	Field Radio Repairman	165
31J2	Teletypewriter Repairman	107
31K2	General Cryptographic Repairman	93
31L2	Field Radio Relay Equipment Repairman	159
31M2	Radio Relay and Carrier Attendant	482

Table A-1 (Continued)

Occupational Area and MOS Course		N[a]
GENERAL ELECTRONIC MAINTENANCE (cont.)		
31Q2	Avionics Repairman (Aviation Electronic Equipment Maintenance)	48 C
31Q3	Avionics Repairman (Aviation Electronic Equipment Repair)	131 C
32B2	Fixed Station Receiver Repairman	184
32E2	Fixed Plant Carrier Repairman	148
32G2	Fixed Cryptographic Equipment Repairman	86
36A1	Wireman	151
36A1	Wireman	149
36A1	Wireman	149
36C2	Lineman	439
36G2	Manual Central Office Repairman	151
36H2	Dial Central Office Repairman	66
PRECISION MAINTENANCE		
41C1	Fire Control Instrument Repair	128
43E2	Parachute Rigger	516
44C2	Welder	148
44E2	Machinist	79
45G2	Turret Artillery Mechanic	196
45G3	Turret Artillery Repairman	155
45J2	Aircraft Armament Repairman	106
46M2	Pershing Rocket Motor/Structures Repairman	11 B

Table A-1 (Continued)

Occupational Area and MOS Course		N[a]
AUXILIARY SERVICES		
51B2	Carpenter	29 B
52B2	Power Generation Equipment Operator/Mechanic	334
54D2	Chemical Equipment Repairman	209
54E2	Chemical Staff Specialist	200
56A1	Supply Handler	58
56C2	Petroleum Storage Specialist	96
56D2	Subsistence Storage Specialist	133
57E2	Laundry, Bath, and Impregnation Specialist	50
MOTORS		
62A1	Engineer Equipment Assistant	99
62B2	Engineer Equipment Repairman	285
62C2	Engineer Missile Equipment Specialist	148
62E2	Construction Machine Operator	214
62F2	Crane-Shovel Operator	19 B
63B2	Wheel Vehicle Repairman	189
63B2	Wheel Vehicle Repairman	201
63B2	Wheel Vehicle Repairman	181
63C2	General Vehicle Repairman	338
63C2	General Vehicle Repairman	302
63F2	Recovery Specialist	170

Table A-1 (Continued)

Occupational Area and MOS Course		N^a
MOTORS (cont.)		
63G2	Fuel and Electrical Systems Repairman	130
63H2	Engine and Power Train Repairman	385
64A1	Light Vehicle Driver	158
64A1	Light Vehicle Driver	135
64A1	Light Vehicle Driver	157
67A1	Aircraft Maintenance Apprentice	360
67D2	Single-Engine Airplane Repairman	166
67Q2	Single-Engine Single-Rotor Helicopter	111
68G2	Airframe Repairman	184
CLERICAL		
71A1	Clerk/Clerk Typist	186
71A1	Clerk/Clerk Typist	141
71A1	Clerk/Clerk Typist	193
71A1	Clerk/Clerk Typist	14 B
71C2	Stenographer	194
71F2	Postal Clerk	58 B
71G2	Medical Records Specialist	237
71H2	Personnel Specialist	109
71H2	Personnel Specialist	93
71H2	Personnel Specialist	56 A
71H2	Personnel Specialist	112
71H2	Personnel Specialist	417

Table A-1 (Continued)

Occupational Area and MOS Course		N[a]
CLERICAL (cont.)		
71N2	Movement Specialist	297
71P2	Flight Operations Coordinator	59 B
72B2	Communications Center Specialist	417
72C2	Telephone Switchboard Operator	335
73A1	Finance Clerk	435
73C2	Pay/Disbursing Specialist	242
73D2	Accounting Specialist	151
74A1	Data Processing Equipment Operator	142
74E2	ADPS Console Operator	61 C
74F2	ADPS Programming Specialist	71 C
76A1	Supplyman	156
76A1	Supplyman	126
76A1	Supplyman	136
76E2	Quartermaster Supply Specialist	238
76J2	Medical Supply and Parts Specialist	152
76K3	General Supply Specialist	109 A
GRAPHICS		
81D2	Map Compiler	69
82C1	Artillery Surveyor	174
82C2	Artillery Survey Specialist	271
82D2	Topographic Surveyor	83
84B2	Still Photographer	115

Table A-1 (Continued)

Occupational Area and MOS Course		N[a]
GENERAL TECHNICAL		
91A1	Med Corpsman	381
91B1	Medical Specialist	175
91D2	Oper Rm Specialist	166
91E2	Dental Specialist	210
91P2	X-Ray Specialist	148
91S1	Preventive Medicine Specialist	131
91R1	Food Inspection Specialist	193
92C2	Petroleum Laboratory Specialist	55
94B2	Cook	109
94B2	Cook	130
94B2	Cook	167
96C2	Interrogator	87 C
96D2	Image Interpreter	46 B
97D2	Military Intelligence Coordinator	125
RADIO CODE		
05B2	Radio Operator	90
05B2	Radio Operator	96 D
05B2	Radio Operator	119
05C2	Radio Teletypewriter Operator	468

[a]Note: The labeled samples were dropped from the analyses for the reason indicated:

A = Low validity
B = Small N
C = Miscellaneous entry MOS
D = No criterion data reported

Table A-2

ACB WEIGHTS AND MULTIPLE CORRELATION COEFFICIENTS[a]
OF ACB TESTS FOR EACH MOS CLUSTER

ACB Test	MOS Cluster IN	AE	EL	GM	MM	CL	GT	RC
VE	-0485	-0207	0481	1105	0458	2639	0780	0772
AR	1548	1346	2263	1608	1262	2583	2285	1632
PA	0326	0307	0775	0409	0084	0216	0719	0598
MA	0376	0476	0661	0842	0531	0003	0779	0480
ACS	0797	2003	0810	0183	0544	1459	1077	0045
ARC	-0197	-0091	0226	0425	0370	0361	0508	1824
SM	1227	0503	0430	1294	-0133	-1156	-0016	0422
AI	0945	2016	0475	1214	2808	0490	-0167	0188
ELI	-1008	-0874	2061	0544	1213	0387	0930	0428
CI	1577	0496	0215	0064	0287	-0013	0272	-0380
GIT	1167	0762	0570	1221	1586	0968	1390	0869
Multiple Correlation	450	469	679	667	668	689	663	583

[a]Decimals omitted.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) U. S. Army Behavioral Science Research Laboratory, Arlington, Virginia	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

DEVELOPMENT OF IMPROVED APTITUDE AREA COMPOSITES FOI ENLISTED CLASSIFICATION

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Milton H. Maier and Edmund F. Fuchs

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
September 1969	51	9

8a. CONTRACT OR GRANT NO. b. PROJECT NO. DA R&D PJ No. 2Q0?210?A722 c. Differential Classification d. b-00	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Research Report 1159 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office, Chief of Research and Development DA, Washington, D. C.

13 ABSTRACT

In response to special requirements of Deputy Chief of Staff for Personnel and United States Continental Army Command, BESRL's DIFFERENTIAL CLASSIFICATION Work Unit has conducted research to develop new and improved aptitude area composites based on the Army Classification Battery (ACB) test scores for use in determining assignment of enlisted input to training courses. Nearing completion is a large-scale validation study of operational and experimental measures conducted across a full range of the Army's Military Occupational Specialties (MOS). The ACB measures were evaluated for effectiveness as predictors of final grades in Army school training courses and performance in Army combat job assignments. Eight new aptitude area composites using only the ACB tests were developed. Simulated allocation studies were conducted to determine how much the new composites would improve classification of men for training and jobs. These composites and benefits to be derived from their use--specifically, enhancement of productivity of enlisted men in training assignments--are described in the present report. With the revised aptitude area system, benefits would accrue to the Army in 1) reduction by about ?% in the number of training school failures; 2) a concomitant savings in training funds and fuller utilization of Army training resources, 3) improvement in training performance of men completing training successfully--the number of men performing at marginal level would decrease by about 1 % and an increased number would perform at superior level. Details of the data collection, statistical analysis and the results are presented in the Technical Supplement section of the report.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
*Differential classification Psychological measurement *Army Classification Battery (ACB) Enlisted input-utilization *Classification tests *Validation techniques Criterion data Allocation studies *Aptitude areas Occupational areas MOS *Sampling techniques Data analysis Statistical analysis *Personnel system--Army						